Justified Belief and the Topology of Evidence

Alexandru Baltag[1], Nick Bezhanishvili[1], Aybüke Özgün[1,2], Sonja Smets[1]

[1] University of Amsterdam, The Netherlands
[2] LORIA, CNRS - Université de Lorraine, France

Abstract. We introduce a new topological semantics for evidence, evidence-based justifications, belief and knowledge. This setting builds on the evidence model framework of van Benthem and Pacuit, as well as our own previous work on (a topological semantics for) Stalnaker's doxastic-epistemic axioms. We prove completeness, decidability and finite model property for the associated logics, and we apply this setting to analyze key issues in Epistemology: "no false lemma" Gettier examples, misleading defeaters, and undefeated justification versus undefeated belief.

1 Introduction

In this paper we propose a topological semantics for *evidence-based belief*, as well as for a notion of "soft" (defeasible) *knowledge*, and explore their connections with various notions of *evidence possession*. This work is largely based on looking from a new perspective at the models for evidence and belief proposed by van Benthem and Pacuit [21], and developed further in [20].

The *basic pieces of evidence* possessed by an agent are modeled as non-empty sets of possible worlds. A *combined evidence* (or just "evidence", for short) is any non-empty intersection of finitely many pieces of evidence. This notion of evidence is not necessarily factive[3], since the pieces of evidence are possibly false (and possibly inconsistent with each other). The family of (combined) evidence sets forms a topological basis, that generates what we call *the evidential topology*. This is the smallest topology in which all the basic pieces of evidence are open, and it will play an important role in our setting. We study the operator of "having (a piece of) evidence for a proposition *P*" proposed by van Benthem and Pacuit, but we also investigate other interesting variants of this concept: "having (combined) evidence for *P*", "having a (piece of) *factive* evidence for *P*" and "having (combined) factive evidence for *P* ". We show that the last notion coincides with the *interior operator* in the evidential topology, thus matching McKinsey and Tarski's original topological semantics for modal logic [15]. We also show that the two factive variants of evidence-possession operators are more expressive than the original (non-factive) one, being able (when interacting with the global modality) to define the non-factive variants, as well as many other doxastic/epistemic operators.

[3] Factive evidence is true in the actual world. In Epistemology it is common to reserve the term "evidence" for factive evidence. But we follow here the more liberal usage of this term in [20], which agrees with the common usage in day to day life, e.g. when talking about "uncertain evidence", "fake evidence", "misleading evidence" etc.

We propose a *'coherentist' semantics for justification and justified belief*, that is obtained by extending, generalizing and (to an extent) "streamlining" the evidence-model framework for beliefs introduced in [21]. An *argument for P* consists of one or more (combined) evidence sets supporting the same proposition *P* (thus providing multiple evidential paths towards a common conclusion). A *justification for P* is an argument for *P* that is consistent with every other evidence. Our proposed definition of belief is equivalent to requiring that: *P is believed iff there is some (evidence-based) justification for P*. According to this setting, in order to believe *P* one needs to have an "undefeated" argument for *P*: one that is not refuted by any available evidence. We show that our notion of belief coincides with the one of van Benthem and Pacuit [21] for *finite* models, but involves a different generalization of their notion in the infinite case. But, in contrast to the later one, our semantics *always ensures consistency of belief*, even when the available pieces of evidence are mutually inconsistent.[4] Our proposal is also very natural from a topological perspective: essentially, *P is believed iff P is true in "almost all" epistemically-possible worlds* (where 'almost all' is interpreted topologically: all except for a nowhere-dense set).

Moving on to 'knowledge', there are a number of different notions one may consider. First, there is "absolutely certain" or "infallible" knowledge, akin to Aumann's concept of 'partitional knowledge' or van Benthem's concept of 'hard information'. In our single-agent setting, this can be simply defined as the global modality (quantifying universally over all epistemically-possible worlds). There are propositions that are 'known' in this infallible way (-e.g. the ones known by introspection or by logical proof), but very few: most facts in science or real-life are unknown in this sense. Hence, it is more interesting to look at notions of knowledge that are less-than-absolutely-certain: so-called 'defeasible knowledge'. The famous Gettier counterexamples [7] show that simply adding "factivity" to belief will not do: true (justified) belief is extremely fragile (i.e. it can be too easily lost), and it is consistent with having only wrong justifications for an (accidentally) true conclusion.

Clark's [5] influential "no false lemma" proposal is to require a *correct justification*: one that doesn't use any falsehood. We formalize this notion by saying that *P is known if there is a factive (true) justification for P*. Note though that our proposal imposes a stronger requirement than Clark's, since our concept of justification requires consistence with all the available (combined) evidence. In our terminology, Clark only requires a factive *argument* for *P*. So Clark's approach is 'local', assessing a knowledge claim based only on the truth of the evidence pieces (and the correctness of the inferences) that are used to justify it. Our proposal is coherentist, and thus 'holistic', assessing knowledge claims by their coherence with all of the agent's acceptance system: justifications need to be checked against all the other arguments that can be constructed from the agent's current evidence.

Another approach to knowledge (also stronger than the no-false-lemma requirement) was championed by Lehrer, Klein and others [13, 14, 11, 12, 17], under the name of

[4] Another, purely technical advantage of our setting is that the resulting doxastic logic has finite model property, in contrast to the one in [21].

"Defeasibility Theory of Knowledge". According to this view, P is known (in the indefeasible sense) only if there is a factive justification for P that cannot be defeated by any further true evidence. This means that the justification is consistent, not only with the currently available evidence, but also with *any potential (new) factive evidence* that the agent might learn in the future. This version of the theory has been criticized as being too strong: some new evidence might be 'misleading' or 'deceiving' despite being true. A weaker version of Defeasibility Theory requires that knowledge is undefeated only by "non-misleading" evidence. In our setting, a proposition P is said to be a *potentially misleading evidence* if it can indirectly generate false evidence (i.e. if by adding P to the family of currently available pieces of evidence we obtain at least one false combined evidence). Misleading propositions include all the false ones, but they may also include some true ones. We show that our notion of knowledge matches this weakened version of Defeasibility Theory (though not the strong version).

Yet another path leading to our setting in this paper goes via our previous work [1, 2] on a topological semantics for the doxastic-epistemic axioms proposed by Stalnaker [18]. These axioms were meant to capture a notion of fallible knowledge, in close interaction with a notion of "strong belief" (defined as "subjective certainty" or the "feeling of knowledge"). The main principle specific to this system was that "believing implies believing that you know" ($Bp \to BKp$), which goes in direct contradiction to Negative Introspection for Knowledge.[5] The topological semantics that we proposed for these concepts in [1, 2] was overly restrictive (being limited to the rather exotic class of "extremally disconnected" topologies). In this paper, we show that these notions can be interpreted on arbitrary topological spaces, without changing their logic. Indeed, our definitions of belief and knowledge above can be seen as the natural generalizations to arbitrary topologies of the notions in [1, 2].

We apply our models to various Gettier-type examples, and completely axiomatize the resulting logics, proving their decidability and finite model property. Our hardest result refers to our richest logic (that can define all the modal operators mentioned above). We end with a discussion of possible research lines for future work.

2 Evidence, Belief and Knowledge in Topological Spaces

2.1 Topological Models for Evidence

Definition 1 (Evidence Models) *(van Benthem and Pacuit)*[6] Given a countable set of propositional letters Prop, an *evidence model* for Prop is a tuple $\mathcal{M} = (X, E_0, V)$, where: X is a non-empty set of "states"; $E_0 \subseteq \mathcal{P}(X) \setminus \{\emptyset\}$ is a family of non-empty sets called *basic evidence sets* (or *pieces of evidence*), with $X \in E_0$; and $V : \text{Prop} \to \mathcal{P}(X)$ is a valuation function.

[5] Indeed, the logic of Stalnaker's knowledge is not $S5$, but the modal logic $S4.2$.

[6] The notion of evidence model in [21] is more general, covering cases in which evidence depends on the actual world, but we stick with what they call 'uniform' models, since this corresponds to restricting to agents who are "evidence-introspective".

Given an evidence model $\mathcal{M} = (X, E_0, V)$, a family $F \subseteq E_0$ of pieces of evidence is *consistent* if $\bigcap F \neq \emptyset$, and *inconsistent* otherwise. A*body of evidence* is a family $F \subseteq E_0$ s.t. every non-empty finite subfamily is consistent. We denote by $\mathcal{F}$ the family of all bodies of evidence, and by $\mathcal{F}^{finite}$ the family of all finite ones. A body of evidence F *supports a proposition* P iff P is true in all worlds satisfying the evidence in F (i.e. $\bigcap F \subseteq P$).

The *strength order* between bodies of evidence is given by inclusion: $F \subseteq F'$ means that F' is at least as strong as F. Note that stronger bodies of evidence support more propositions: if $F \subseteq F'$ then every proposition supported by F is also supported by F'. A body of evidence is *maximal* ("strongest") if it's not included in any other such body. We denote by $Max_{\subseteq}\mathcal{F} = \{F \in \mathcal{F} : \forall F' \in \mathcal{F}(F \subseteq F' \Rightarrow F = F')\}$ the family of all maximal bodies of evidence. By Zorn's Lemma, *every body of evidence can be strengthened to a maximal body of evidence*: $\forall F \in \mathcal{F} \exists F' \in Max_{\subseteq}\mathcal{F}(F \subseteq F')$.

A *combined evidence* (or just "*evidence*", for short) is any non-empty intersection of finitely many pieces of evidence. We denote by $E := \{\bigcap F : F \in \mathcal{F}^{finite} \text{ s.t. } \bigcap F \neq \emptyset\}$ the family of all (combined) evidence.[7] A (combined) evidence $e \in E$ *supports* a proposition $P \subseteq X$ if $e \subseteq P$. (In this case, we also say that e is *evidence for* P.) Note that the natural strength order between combined evidence sets goes the other way around (reverse inclusion): $e \supseteq e'$ means that e' is at least as strong as e.[8]

The intuition is that $e \in E_0$ represent the basic pieces of "direct" evidence (obtained say by observation or via testimony) that are possessed by the agent, while the combined evidence $e \in E$ represents indirect evidence that is obtained by combining finitely many pieces of direct evidence. Not all of this evidence is necessarily true though.

We say that some (basic or combined) evidence $e \in E$ is *factive evidence* at world $x \in X$ whenever it is true at x (i.e. $x \in e$). A body of evidence F is factive if all the pieces of evidence in F are factive (i.e. $x \in \bigcap F$).

The *plausibility (pre)order* $\sqsubseteq_E$ associated to an evidence model is given by:

$$x \sqsubseteq_E y \text{ iff } \forall e \in E_0\,(x \in e \Rightarrow y \in e) \text{ iff } \forall e \in E\,(x \in e \Rightarrow y \in e).$$

Definition 2 (Topological Space) A *topological space* is a pair $\mathcal{X} = (X, \tau)$, where X is a non-empty set and τ is a *topology* on X, i.e. a family $\tau \subseteq \mathcal{P}(X)$ containing X and $\emptyset$, and closed under finite intersections and arbitrary unions. Given a family $E \subseteq \mathcal{P}(X)$ of subsets of X, the *topology generated by* E is the smallest topology τ_E on X such that $E \subseteq \tau_E$. A set $A \subseteq X$ is *closed* iff it is the complement of an open set, i.e. it is of the form $X \setminus U$ with $U \in \tau$. Let $\tau^c = \{X \setminus U \,|\, U \in \tau\}$ denote the family of all closed sets of $\mathcal{X} = (X, \tau)$. In any topological space $\mathcal{X} = (X, \tau)$, one can define two important operators, namely *interior* $Int : \mathcal{P}(X) \to \mathcal{P}(X)$ and *closure* $Cl : \mathcal{P}(X) \to \mathcal{P}(X)$, given by

[7] This is a difference in notation with the setting in [21, 20], where E is used to denote the family of basic evidence sets (denoted here by E_0).

[8] This is both to fit with the strength order on bodies of evidence (since $F \subseteq F'$ implies $\bigcap F \supseteq \bigcap F'$), and to ensure that stronger evidence supports more propositions: since, if $e \supseteq e'$, then every proposition supported by e is supported by e'.

$IntP := \bigcup\{U \in \tau \mid U \subseteq P\}$, $ClP := \bigcap\{C \in \tau^c \mid P \subseteq C\}$. A set $A \subseteq X$ is called *dense* in $\mathcal{X}$ if $ClA = X$ and it is called *nowhere dense* if $IntClA = \emptyset$. For a topological space $\mathcal{X} = (X, \tau)$, the *specialization preorder* $\sqsubseteq_\tau$ is given by: $x \sqsubseteq_\tau y$ iff $\forall U \in \tau\, (x \in U \Rightarrow y \in U)$.

Special Case: Relational Spaces. A topological space is called *Alexandroff* iff the topology is closed under arbitrary intersections. An Alexandroff topology is fully captured by its specialization preorder: *in this case, the interior operator coincides with the Kripke modality for the specialization relation* (i.e. $IntP = \{x \in X \mid \forall y\, (x \sqsubseteq_\tau y \Rightarrow y \in P)\}$). There is a canonical bijection between Alexandroff topologies $\mathcal{X} = (X, \tau)$ and *preordered spaces*[9] $(X, \leq)$, mapping (X, τ) to $(X, \sqsubseteq_\tau)$; the inverse map takes $(X, \leq)$ into $(X, Up(X))$, where $Up(X)$ is the family of upward-closed sets[10].

An Even More Special Case: (Grove/Lewis) Sphere Spaces. These are topological spaces in which the opens are "nested", i.e. for every $U, U' \in \tau$, we have either $U \subseteq U'$ or $U' \subseteq U$. Sphere spaces are Alexandroff, and moreover they correspond exactly to *totally preordered spaces* (i.e. sets X endowed with a total preorder $\leq$).

Definition 3 (Topological Evidence Models) A *topological evidence model* ("*topo-e-model*", for short) is a structure $\mathcal{M} = (X, E_0, \tau, V)$, where (X, E_0, V) is an evidence model and $\tau = \tau_E$ is the topology generated by the family of combined evidence E (or equivalently, by the family of basic evidence sets E_0)[11], which will be called *the evidential topology*. It is easy to see that *the plausibility order $\sqsubseteq_E$ of $\mathcal{M}$ coincides with the specialization order* of the associated topology: $\sqsubseteq_E = \sqsubseteq_\tau$.

Since any family $E_0 \subseteq \mathcal{P}(X)$ generates a topology, topo-e-models are just another presentation of (uniform) evidence models. We use this special terminology to stress our focus on the topology, and to avoid ambiguities (since our definition of belief in topo-e-models will be different from the definition of belief in evidence models in [21]).

A topo-e-model is said to be *Alexandroff* iff the underlying topology is Alexandroff. So they can be understood as *relational (plausibility) models*, in terms of a preorder $\leq$ ("plausibility relation"). A special case is the one of *Grove-Lewis (topological) evidence models*: this is the case when the basic pieces of evidence are nested (i.e. for all $e, e' \in E_0$ we have either $e \subseteq e'$ or $e' \subseteq e$). It is easy to see that in this case all the opens of the generated topology are also nested, so the topology is that of a sphere space.

Proposition 1 *Given a topo-e-model $\mathcal{M} = (X, E_0, \tau, V)$, the following are equivalent:*

1. *$\mathcal{M}$ is Alexandroff;*
2. *The family E of (combined) evidence is closed under arbitrary non-empty intersections (i.e. if $F \subseteq E$ and $\bigcap F \neq \emptyset$, then $\bigcap F \in E$);*
3. *Every consistent body of evidence is equivalent to a finite body of evidence (i.e. $\forall F \in \mathcal{F} (\bigcap F \neq \emptyset \Rightarrow \exists F' \in \mathcal{F}^{finite}$ s. t. $\bigcap F = \bigcap F'$)).*

---

[9] A *preorder* on X is a reflexive-transitive relation on X.
[10] A subset $A \subseteq X$ is said to be *upward-closed* wrt $\leq$ if $\forall x, y \in X\, (x \in A \wedge x \leq y \Rightarrow y \in A)$.
[11] These families generate the same topology. We denote it by τ_E only because the family E of combined evidence forms a *basis* of this topology.

Arguments and Justifications. We can use this setting to formalize a "coherentist" view on justification. An *argument for* P is a disjunction $U = \bigcup_{i \in I} e_i$ of (some non-empty family of) (combined) evidences $e_i \in E$ that all support P (i.e. $e_i \subseteq P$ for all $i \in I$). Thus, an argument may provide *multiple evidential paths* e_i to support a common conclusion P. Topologically, an argument for P is the same as a *non-empty open subset of* P ($U \in \tau_E$ s.t. $U \subseteq P$). Also, the interior $IntP$ is the *weakest (most general) argument for* P.

A *justification for* P is an argument U for P that is consistent with every (combined) evidence (i.e. $U \cap e \neq \emptyset$ for all $e \in E$, which in fact implies that $U \cap U' \neq \emptyset$ for all $U' \in \tau_E \setminus \{\emptyset\}$). So justifications are arguments that are not defeated by any available evidence. Topologically, we can see that a justification for P is just an *(everywhere) dense open subset* of P (i.e. $U \in \tau_E$ s.t. $U \subseteq P$ and $Cl_{\tau_E}(U) = X$). As for evidence, an argument or a justification for P is said to be *factive* (or "correct") if it is true in the actual world. The fact that arguments are open in the generated topology encodes the principle that *any argument should be evidence-based*: whenever an argument is correct, then it is supported by some factive evidence. To anticipate further: in our setting, justifications will form the basis of *belief*, while correct justifications will form the basis of *(defeasible) knowledge*. But for now we'll introduce a stronger form of "knowledge": the absolutely-certain and irrevocable kind.

Infallible Knowledge: possessing hard information. We use $\forall$ for the so-called *global* modality, which associates to every proposition $P \subseteq X$, some other proposition $\forall P$, given by putting: $(\forall P) := X$ iff $P = X$, and $(\forall P) := \emptyset$ otherwise. In other words: $(\forall P)$ holds (at any state) iff P holds at all states. In this setting, $\forall$ is interpreted as "absolutely certain, *infallible knowledge*", defined as truth in all the worlds that are consistent with the agent's information.[12] This is not a realistic concept of knowledge, but just a limit notion, encompassing all epistemic possibilities.

Having *Basic* Evidence for a Proposition. van Benthem and Pacuit define, for every proposition $P \subseteq X$, another proposition[13] $E_0 P$ given by putting: $E_0 P := X$ if $\exists e \in E_0\,(e \subseteq P)$, and $E_0 P := \emptyset$ otherwise. Essentially, $E_0 P$ means that "the agent *has basic evidence for* P", i.e. P is supported by some available piece of evidence. One can also introduce a *factive* version of this proposition: $\Box_0 P$, read as "the agent has *factive basic evidence for* P", is given by putting

$$\Box_0 P := \{x \in X : \exists e \in E_0\,(x \in e \subseteq P)\}.$$

Having (Combined) Evidence for a Proposition. If in the above definitions of $E_0 P$ and $\Box_0 P$ we replace basic pieces of evidence by combined evidence, we obtain two other operators EP, meaning that "the agent *has (combined) evidence for* P", and $\Box P$,

[12] In a multi-agent model, some worlds might be consistent with one agent's information, while being ruled out by another agent's information. So, in a multi-agent setting, $\forall_i$ will only quantify over all the states in agent i's current information cell (according to a partition Π_i of the state space reflecting agent i's hard information).

[13] They denote this by EP, but we use $E_0 P$ for this notion, since we reserve the notation EP for having *combined* evidence for P.

meaning that "the agent *has factive (combined) evidence for P*". More precisely:

$$EP := X \text{ if } \exists e \in E\, (e \subseteq P), \quad \text{and } EP := \emptyset \text{ otherwise;}$$

$$\Box P := \{x \in X : \exists e \in E\, (x \in e \subseteq P)\}.$$

Observation 1. Note that *the agent has evidence for a proposition P iff she has an argument for P*. So EP can also be interpreted as "having an argument for P". Similarly, $\Box P$ can be interpreted as "having a *correct (i.e. factive) argument* for P".

Observation 2. Note that the agent has factive evidence for P at x iff x is in the interior of P. So our modality $\Box$ *coincides with the interior operator*: $\Box P = IntP$.

2.2 Belief

Belief à la van Benthem-Pacuit [21]. The notion of belief proposed by van Benthem and Pacuit, which we will denote by *Bel*, is that *P is believed iff every maximal body of evidence supports P*: $BelP$ holds (at any state of X) iff we have $\bigcap F \subseteq P$ for every $F \in Max_{\subseteq}\mathcal{F}$. As already noticed in [21], this is equivalent to treating evidence models as special cases of plausibility models [3, 4, 19], with the plausibility relation given by $\sqsubseteq_E$ (or equivalently, as Grove-Lewis "sphere models" [9] where the spheres are the sets that are upward closed wrt $\sqsubseteq_E$), and applying the standard definition (due to Grove) of belief as "truth in all the most plausible worlds".[14] Grove's definition works well when the plausibility relation is well-founded (and also in the somewhat more general case given by the Grove-Lewis Limit Assumption), but it yields inconsistent beliefs in the case that there are *no* most plausible worlds. But note that in evidence models $\sqsubseteq_E$ may be non-wellfounded. Indeed, *belief à la van Benthem-Pacuit can be inconsistent*:

Example 1 *Consider the evidence model $\mathcal{M} = (\mathbb{N}, E_0, V)$, where the state space is the set $\mathbb{N}$ of natural numbers, $V(p) = \emptyset$, and the basic evidence family $E_0 = \{e \subseteq \mathbb{N} : \mathbb{N} \setminus e \text{ finite}\}$ consists of all co-finite sets. The only maximal body of evidence in E_0 is E_0 itself. However, $\bigcap E_0 = \emptyset$. So $Bel\bot$ holds in $\mathcal{M}$.*

This phenomenon only happens in (some cases of) *infinite* models, so it is *not* due to the inherent mutual inconsistency of the available evidence. The "good" examples in [21] are the ones in which (possibly inconsistent) evidence is processed to yield consistent beliefs. So it seems to us that the intended goal (only partially fulfilled) in [21] was to ensure that the agents are able to form *consistent beliefs* based on the available evidence. We think this to be a natural requirement for idealized "rational" agents, and so we consider doxastic inconsistency to be "a bug, not a feature", of the van Benthem-Pacuit framework. Hence, we now propose a notion that agrees with the one in [21] in all the "good" cases, but also produces in a natural way only consistent beliefs.

[14] Note that all the notions of belief we consider are global: they do not depend on the state of the world, i.e. we have either $BelP = X$ or $BelP = \emptyset$ (similar to the sets $\forall P, E_0P, EP$). This expresses the assumption that belief is a purely internal notion, thus transparent and hence absolutely introspective. This is standard in logic and accepted by most philosophers.

Our Notion of Belief. The intuition is that P is believed iff *it is entailed by all the "sufficiently strong" (combined) evidence*. Formally, BP holds iff *every finite body of evidence can be strengthened to some finite body of evidence which supports P*:

$$BP \text{ holds (at any state)} \quad \text{iff} \quad \forall F \in \mathcal{F}^{finite} \exists F' \in \mathcal{F}^{finite} (F \subseteq F' \wedge \bigcap F' \subseteq P).$$

Our notion of belief B coincides with *Bel* in the *finite* case, or, more generally, in *evidence models in which every maximal body of evidence is consistent*. But, unlike *Bel*, our notion of belief B is *always consistent* (i.e. $B\bot = B\emptyset = \emptyset$), and moreover it satisfies the axioms of the standard doxastic logic $KD45$. Another nice feature is that our belief B is a *purely topological notion*, as can be seen from the following:

Proposition 2 *In every evidence model* (X, E_0, V)*, the following are equivalent, for any proposition* $P \subseteq X$*:*

1. *BP holds (at any state);*
2. *every (combined) evidence can be strengthened to some evidence supporting P* ($\forall e \in E \exists e' \in E$ *s.t* $e' \subseteq e \cap P$*);*
3. *every argument (for anything) can be strengthened to an argument for P* ($\forall U \in \tau_E \setminus \{\emptyset\} \exists U' \in \tau_E \setminus \{\emptyset\}$ *s.t.* $U' \subseteq U \cap P$*);*
4. there is a justification for *P: i.e. some argument for P which is consistent with any available evidence* ($\exists U \in \tau_E$ *s.t.* $U \subseteq P$ *and* $U \cap e \neq \emptyset$ *for all* $e \in E$*);*
5. *P includes some dense open set;*
6. *IntP is* dense *in* τ_E *(i.e.* $Cl(IntP) = X$*), or equivalently* $X \setminus P$ is nowhere dense*;*
7. $\forall\Diamond\Box P$ *holds (at any state: i.e.* $\forall\Diamond\Box P \neq \emptyset$*, or equivalently* $\forall\Diamond\Box P = X$*), where* $\Diamond P := \neg\Box\neg P$ *is the dual of the* $\Box$ *operator.*

Proposition 2 part (4) can be interpreted as saying that *our notion of belief B is the same as "justified belief"*: a proposition P is believed iff the agent has a justification for P. In this case, there exists a *weakest (most general) justification* for P, namely *IntP*. Moreover, part (6) shows that our proposal is very natural from a topological perspective: it is equivalent to saying that *P is believed iff the complement of P is nowhere dense*. Since nowhere dense sets are one of the topological concepts of "small" or "negligible sets", this amounts to *believing propositions if they are true in "almost all" epistemically-possible worlds* (where 'almost all' is interpreted topologically). Finally, part (7) tells us that *belief is definable in terms of the operators* $\forall$ *and* $\Box$.

Our notion of belief can be viewed as a formalization of a "coherentist" epistemology of belief. The requirement that a belief's justification must be *open* in the evidential topology simply means that the justification is ultimately based on the available evidence; while the requirement that the justification is *dense* (in the same topology) means that all the agent's beliefs must be coherent with all her evidence.[15]

[15] Lehrer uses the metaphor of a Subjective Justification Game [13]: rational beliefs are based on justifications that survive a game between the Believer and an inner Critic, who tries to defeat them using the Believer's own "acceptance system".

Conditional Belief. For sets $Q, Q' \subseteq X$, we say that Q' is *Q-consistent* iff $Q \cap Q' \neq \emptyset$. A body of evidence F is *Q-consistent* iff $\bigcap F \cap Q \neq \emptyset$. We say that *$P$ is believed given Q*, and write $B^Q P$, iff *every finite Q-consistent body of evidence can be strengthened to some finite Q-consistent body of evidence supporting $Q \to P$* (i.e. $\neg Q \cup P$). Similarly to Proposition 2, $B^Q P$ is equivalent to any of the following: every Q-consistent evidence can be strengthened to some Q-consistent evidence supporting $Q \to P$; every Q-consistent argument can be strengthened to a Q-consistent argument for $Q \to P$; there is a Q-consistent argument for $Q \to P$ which is consistent with any Q-consistent evidence; $Q \to P$ includes some Q-consistent open set which is dense in Q; $\forall(Q \to \Diamond(Q \wedge \Box(Q \to P))) = X$; etc.

2.3 Knowledge

We now define a "softer" notion of knowledge, that is closer to the common usage of the word than "infallible" knowledge. Formally, we put $KP := \{x \in X : \exists U \in \tau(x \in U \subseteq P \wedge Cl(U) = X)\}$. So KP holds at x iff *P includes a dense open neighborhood of x*; equivalently, iff $x \in IntP$ and $IntP$ is dense. Essentially, this says that *knowledge is "correctly justified belief"*: KP holds at world x iff there exists some justification $U \in \tau$ for P such that $x \in U$. In other words, P is known iff there *exists some correct (i.e. factive) argument for P that is consistent with all the available evidence*.

Note that K satisfies Stalnaker's Strong Belief Principle $BP = BKP$: *from a subjective point of view, belief is indistinguishable from knowledge* [18].[16]

Example 2 *Consider the model $\mathcal{X} = ([0,1], E_0, V)$, where $E_0 = \{(a,b) \cap [0,1] : a, b \in \mathbb{R},\ a < b\}$ and $V(p) = \emptyset$. The generated topology τ_E is the standard topology on $[0,1]$. Let $P = [0,1] \setminus \{\frac{1}{n} : n \in \mathbb{N}\}$ be the proposition stating that the actual state is* not *of the form $\frac{1}{n}$, for any $n \in \mathbb{N}$. Since the complement $\neg P = \{\frac{1}{n} : n \in \mathbb{N}\}$ is nowhere dense, the agent* believes *P, and e.g. $U = \bigcup_{n \geq 1}(\frac{1}{n+1}, \frac{1}{n})$ is a (dense, open) justification for P. This belief is* true at world $0 \in P$. *But this* true belief is not knowledge *at 0: no justification for P is true at 0, since P doesn't include any open neighborhood of 0, so $0 \notin IntP$ and hence $0 \notin KP$. (However, P is known at all the* other *worlds $x \in P \setminus \{0\}$, since $\forall x \in P \setminus \{0\} \exists \epsilon > 0$ s.t. $x \in (x - \epsilon, x + \epsilon) \subseteq P$, hence $x \in IntP$.)*

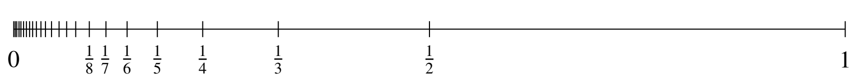

Fig. 1. $([0,1], \tau_E)$

[16] As we'll see, K and B satisfy all the Stalnaker axioms for knowledge and belief [1, 2, 16] and further generalizes our previous work on a topological interpretation of Stalnaker's doxastic-epistemic axioms, which was based on extremally disconnected spaces.

This 'soft' type of knowledge is *defeasible*. In contrast, the usual assumption in Logic is that *knowledge acquisition is monotonic*. As a result, logicians typically assume that knowledge is "irrevocable": once acquired, it cannot be defeated by any further evidence. In our setting, the only irrevocable knowledge is the infallible one, captured by the operator $\forall$. Clearly, K is *not* irrevocable.

Epistemologists have made various other proposals on how a realistic concept of knowledge should be defined. A conception that is very close to (though subtly different from) our notion is the one held by the proponents of the so-called Defeasibility Theory of Knowledge, e.g. Lehrer and Paxson [14], Lehrer [13], Klein [11, 12]: "in-defeasible knowledge" cannot be defeated by any *factive* evidence that might be gathered later (though it may be defeated by false "evidence"). In its simplest version, this says that "an agent knows that P if and only if P is true, she believes that P, and she continues to believe P if any *true* information is received" (Stalnaker [18]). In our formalism, this would require P to be believed conditional on every true "new evidence": i.e. P is known in world x iff $B^Q P$ holds for every $Q \subseteq X$ with $x \in Q$. This simple version is what Rott calls "the Stability Theory of Knowledge" [17]. In contrast, the full-fledged version of the Defeasibility Theory, as held by Lehrer and others, insists that, in order to know P, not only the belief in P has to stay undefeated, but also its justification (i.e. what we call here "an argument for P"). In other words, there must exist an argument for P that is believed conditional on every true evidence. Clearly, this implies that the belief in P is stable; but the converse is not at all obvious. Indeed, Lehrer claims that the converse is false. The problem is that, when confronted with various new pieces of evidence, the agent might keep switching between different justifications (for believing P); thus, she may keep believing in P conditional on any such new true evidence, without actually having any "good" justification (i.e. one that remains itself undefeated by all true evidence). To have 'knowledge', we thus need a *stable justification*.[17]

However, many authors attacked the above interpretation (of both the stability and the defeasibility theory) as being *too strong*: if we allow as potential defeaters *all* factive propositions (i.e. all sets of worlds P containing the actual world), then there are intuitive examples showing that knowledge KP *can* be defeated. Here is such an example, discussed by a leading proponent of the defeasibility theory (Klein [12]). Loretta filled in her federal taxes, following very carefully all the required procedures on the forms, doing all the calculations and double checking everything. Based on this evidence, she correctly believes that she owes $500, and she seems perfectly justified to believe this. So it seems obvious that she *knows* this. But suppose now that, being aware of her own fallibility, she asks her accountant to check her return. The accountant finds no errors, and so he sends her his reply reading "Your return contains *no* errors"; but he inadvertently leaves out the word "no". If Loretta would learn the true fact that the accountant's letter actually reads "Your return contains errors", she would lose her belief that she owed $500! So it seems that there exist defeaters that are true but "misleading".

---

[17] Lehrer uses the metaphor of an 'Ultra-Justification Game' [13], according to which 'knowledge' is based on arguments that survive a game between the Believer and an omniscient truth-telling Critic, who tries to defeat the argument by using both the Believer's current "justification system" *and any new true evidence*.

We can formalize this counterexample as follows.

Example 3 *Consider the model* $\mathcal{M} = (X, E_0, V)$*, where* $X = \{x_1, x_2, x_3, x_4, x_5\}$*,* $V(p) = \emptyset$*,* $E_0 = \{X, O_1, O_2\}$*,* $O_1 = \{x_1, x_2, x_3\}$*,* $O_2 = \{x_3, x_4, x_5\}$*. The resulting set of combined evidence is* $E = \{X, O_1, O_2, \{x_3\}\}$*. Assume the actual world is* x_1*. Then* O_1 *is* known*, since* $x_1 \in Int(O_1) = O_1$ *and* $Cl(O_1) = X$*. Now consider the model* $\mathcal{M}^{+O_3} = (X, E_0^{+O_3}, V)$ *obtained by adding the new evidence* $O_3 = \{x_1, x_5\}$*. We have* $E_0^{+O_3} = \{X, O_1, O_2, O_3\}$*, so* $E^{+O_3} = \{X, O_1, O_2, O_3, \{x_1\}, \{x_3\}, \{x_5\}\}$*. Note that the new evidence is* true *(*$x_1 \in O_3$*). But* O_1 *is* not even believed *in* $\mathcal{M}^{+O_3}$ *anymore (since* $O_1 \cap \{x_5\} = \emptyset$*, so* O_1 *is no longer dense in* $\tau_{E^{+O_3}}$*), thus* O_1 *is* no longer known *after the true evidence* O_3 *was added!*

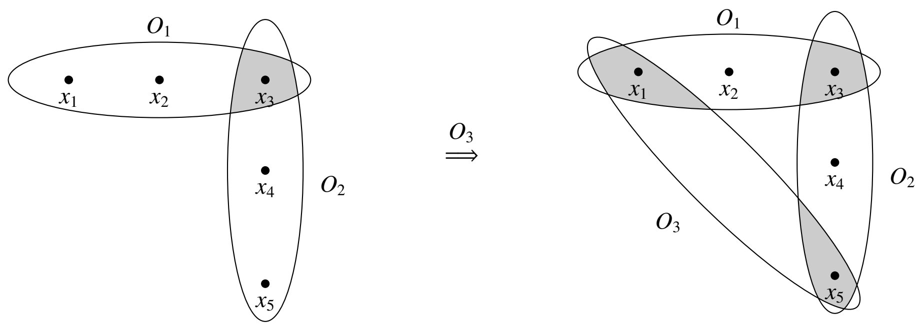

Fig. 2. From $\mathcal{M}$ to $\mathcal{M}^{+O_3}$

Klein's story corresponds to taking O_1 to represent Loretta's direct evidence (based on careful calculations) that she owes \$500, O_2 to represent her prior evidence (based on past experience) that the accountant doesn't make mistakes in his replies to her, and O_3 the potential new evidence provided by the letter. In conclusion, our notion of knowledge is incompatible with the above-mentioned strong interpretations of both stability and defeasibility theory, thus confirming the objections raised against them.

Klein's solution is that one should exclude such 'misleading' defeaters, which may "unfairly" defeat a good justification. But how can we distinguish them from genuine defeaters? Klein's diagnosis, in Foley's more succinct formulation [6], is that "a defeater is misleading if it justifies a falsehood in the process of defeating the justification for the target belief". In the example, the falsehood is that the accountant had discovered errors in Loretta's tax return. It seems that the new evidence O_3 (the existence of the letter as actually written) supports this falsehood, but how? According to us, it is the combination $O_2 \cap O_3$ of the new (true) evidence O_3 with the old (false) evidence O_2 that supports the new falsehood: the true fact (about the letter saying what it says) entails a falsehood *only* if it is taken in conjunction with Loretta's prior evidence (or blind trust) that the accountant cannot make mistakes. So intuitively, *misleading defeaters are the ones which may lead to new false conclusions when combined with some of the old evidence*.

We proceed now to formalize this distinction. Given a topo-e-model $\mathcal{M}$, a proposition $Q \subseteq X$ is *misleading at* $x \in X$ *wrt* E if evidence-addition with Q produces some false new evidence; i.e. if there is some $e' \in E^{+Q} \setminus E$ s.t. $x \notin e'$; equivalently, there is some $e \in E$ s.t. $x \notin (e \cap Q) \notin E \cup \{\emptyset\}$. It is easy to see that: *old evidence in E is by definition non-misleading wrt E* (i.e. if $e \in E$ then e is non-misleading wrt E), and *new non-misleading evidence must be true* (i.e. if $Q \notin E$ is non-misleading at x then $x \in Q$).

We are now in the position to formulate precisely the "weakened" versions of both stability and defeasibility theory that we are looking for. The Weak Stability Theory will stipulate that P is known if it is undefeated by every non-misleading proposition: i.e. $B^Q P$ holds for every non-misleading $Q \subseteq X$. The Weak Defeasibility Theory will require that there exists some justification (argument) for P that is undefeated by every non-misleading proposition. Finally, there is a third formulation, which one might call Epistemic Coherence theory, saying that P is known iff there exists some justification (argument) for P which is consistent with every non-misleading proposition.

The following counterexample shows that weak stability is (only a necessary, but) not a sufficient condition for knowledge:

Example 4 *Consider the model* $\mathcal{M} = (X, E_0, V)$*, where* $X = \{x_0, x_1, x_2\}$*,* $V(p) = \emptyset$*,* $E_0 = \{X, O_1, O_2\}$*,* $O_1 = \{x_1\}$*,* $O_2 = \{x_1, x_2\}$*. The resulting set of combined evidence is* $E = E_0$*. Assume the actual world is* x_0*, and let* $P = \{x_0, x_1\}$*. Then* P *is* believed *(since its interior* $IntP = \{x_1\}$ *is dense) but it is* not known *(since* $x_0 \notin IntP = \{x_1\}$*). However, we can show that* P is believed conditional on any non-misleading proposition. *For this, note that the family of non-misleading propositions (at* x_0*) is* $E \cup \{P, \{x_0\}\} = \{X, O_1, O_2, P, \{x_0\}\}$*. It is easy to see that for each set* Q *in this family, we have* $B^Q P$*.*

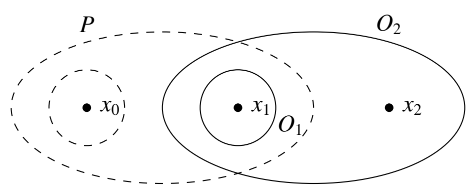

Fig. 3. $\mathcal{M} = (X, E_0, V)$: The continuous ellipses represent the currently available pieces of evidence, while the dashed ones represent the other non-misleading propositions.

One should stress that our counterexample agrees with the position taken by most proponents of Defeasibility Theory: stability of (justified) belief is not enough for knowledge. Intuitively, what happens in the above example is that, although the agent continues to believe P given any non-misleading evidence, her justification keeps changing: there is *no* uniform justification for P that works for every non-misleading evidence Q.

The next result shows that *our notion of knowledge exactly matches the weakened version of Defeasibility Theory*, as well as the Epistemic Coherence formulation:

Proposition 3 *Let $\mathcal{M}$ be a topo-e-model, and assume $x \in X$ is the actual world. The following are equivalent for all $P \subseteq X$:*

1. *P is known ($x \in KP$).*
2. *there is an argument for P that cannot be defeated by any non-misleading proposition; i.e. $\exists U \in \tau_E \setminus \{\emptyset\}$ s.t. $U \subseteq P$ and $B^Q U$ for all non-misleading $Q \subseteq X$.*
3. *there is an argument for P that is consistent with every non-misleading proposition; i.e. $\exists U \in \tau_E \setminus \{\emptyset\}$ s.t. $U \subseteq P$ and $U \cap Q \neq \emptyset$ for all non-misleading $Q \subseteq X$.*

3 Logics for evidence, belief and knowledge

In this section, we present formal languages for evidence, belief and knowledge, and provide sound, complete and decidable proof systems for the resulting logics.

The *topological language* $\mathcal{L}$ is given by the following grammar

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid B\varphi \mid K\varphi \mid \forall\varphi \mid B^{\varphi}\varphi \mid \Box\varphi \mid E\varphi$$

where $p \in$ Prop. We employ the usual abbreviations for propositional connectives $\top$, $\bot$, $\vee$, $\rightarrow$, $\leftrightarrow$, and for the dual modalities $\langle B\rangle$, $\langle K\rangle$, $\langle E\rangle$ etc, except that some of them have special abbreviations: $\exists\varphi := \langle\forall\rangle\varphi$ and $\Diamond\varphi := \langle\Box\rangle\varphi$.

Several fragments of $\mathcal{L}$ have special importance: $\mathcal{L}_B$ is the fragment having the belief B as the only modality; $\mathcal{L}_K$ has only the knowledge operator K; $\mathcal{L}_{KB}$ has only operators K and B; $\mathcal{L}_{\forall K}$ has only operators $\forall$ and K; $\mathcal{L}_{\forall\Box}$ has only operators $\forall$ and $\Box$.

We also consider an *extension* $\mathcal{L}_{E_0\Box_0}$ of $\mathcal{L}$, called the *evidence language*: this is obtained by extending $\mathcal{L}$ with *two new operators E_0 and $\Box_0$*. The expressivity of $\mathcal{L}_{E_0\Box_0}$ goes beyond purely topological properties: the meaning of E_0 and $\Box_0$ does not depend only on the topology, but also on the basic evidence family E_0. Finally, we will consider one very important fragment of $\mathcal{L}_{E_0\Box_0}$, namely the language $\mathcal{L}_{\forall\Box\Box_0}$ having only the operators $\forall$, $\Box$ and $\Box_0$. Its importance comes from that $\mathcal{L}_{\forall\Box\Box_0}$ is *co-expressive* with $\mathcal{L}_{E_0\Box_0}$.

The *semantics* for these languages is obvious: given a topo-e-model $\mathcal{M} = (X, E_0, \tau, V)$, we recursively extend the valuation map V to an interpretation map $||\varphi||$ for all formulas φ, by interpreting the Boolean connectives and the modalities using the corresponding semantic operators: e.g. $||\forall\varphi|| = \forall||\varphi||$, $||\Box\varphi|| = \Box||\varphi||$ etc.

Proposition 4 *The following equivalences are valid in all topo-e-models:*

1. $B\varphi \leftrightarrow \langle K\rangle K\varphi \leftrightarrow \exists K\varphi \leftrightarrow \forall\Diamond\Box\varphi$	4. $K\varphi \leftrightarrow \Box\varphi \wedge B\varphi \leftrightarrow \Box\varphi \wedge \forall\Diamond\Box\varphi$
2. $E\varphi \leftrightarrow \exists\Box\varphi$	5. $B^{\theta}\varphi \leftrightarrow \forall(\theta \rightarrow \Diamond(\theta \wedge \Box(\theta \rightarrow \varphi)))$
3. $E_0\varphi \leftrightarrow \exists\Box_0\varphi$	6. $\forall\varphi \leftrightarrow B^{\neg\varphi}\bot$

So, all the other modalities of $\mathcal{L}_{E_0\Box_0}$ can be defined in $\mathcal{L}_{\forall\Box\Box_0}$.

Theorem 1 *The system $KD45$ (for the B operator) is sound and complete for $\mathcal{L}_B$.*

Theorem 2 *The system $S4.2$ (for the K operator) is sound and complete for $\mathcal{L}_K$.*

Theorem 3 *A sound and complete axiomatization for $\mathcal{L}_{KB}$ is given by Stalnaker's system* [18] *KB in [18], consisting of the following:*

1. *the $S4$ axioms and rules for Knowledge K*
2. *Consistency of Belief: $B\phi \to \neg B\neg\phi$;*
3. *Knowledge implies Belief: $K\phi \to B\phi$;*
4. *Strong Positive and Negative Introspection for Belief: $B\phi \to KB\phi$; $\neg B\phi \to K\neg B\phi$;*
5. *the "Strong Belief" axiom: $B\phi \to BK\phi$.*

Theorem 4 ([8]) *The following system is sound and complete for $\mathcal{L}_{\forall\Box}$:*

1. *the S5 axioms and rules for $\forall$*
2. *the S4 axioms and rules for $\Box$*
3. $\forall\varphi \to \Box\varphi$

By Proposition 4, $\mathcal{L}_{\forall\Box}$ can define *all* the other operators of $\mathcal{L}$. So a complete system for $\mathcal{L}$ is obtained by adding the relevant axiom-definitions to the above system.

Theorem 5 *The following system is sound and complete for $\mathcal{L}_{\forall K}$:*

1. *the $S5$ axioms and rules for $\forall$*
2. *the $S4$ axioms and rules for K*
3. $\forall\varphi \to K\varphi$
4. $\exists K\varphi \to \forall\langle K\rangle\varphi$

Since belief is definable in $\mathcal{L}_{\forall K}$, a complete system for the language with this additional belief operator is obtained by adding the axiom-definition $B\varphi \leftrightarrow \exists K\varphi$ to the above system for $\mathcal{L}_{\forall K}$.

Theorem 6 (Soundness, Completeness, Finite Model Property and Decidability) *The logic $\mathcal{L}_{\forall\Box\Box_0}$ is* completely axiomatizable *and has the* finite model property, *and hence it is decidable. A complete axiomatization is given by the following system $L_{\forall\Box\Box_0}$:*

1. *the S5 axioms and rules for $\forall$*
2. *the S4 axioms and rules for $\Box$*
3. $\Box_0\varphi \to \Box_0\Box_0\varphi$
4. *the Monotonicity Rule for $\Box_0$: from $\varphi \to \psi$, infer $\Box_0\varphi \to \Box_0\psi$*
5. $\forall\varphi \to \Box_0\varphi$
6. $\Box_0\varphi \to \Box\varphi$
7. *the Pullout Axiom*[19]*:* $(\Box_0\varphi \wedge \forall\psi) \to \Box_0(\varphi \wedge \forall\psi)$

[18] This shows that the semantics in this paper correctly generalizes the one in [1, 2, 16] for the system KB.

[19] This axiom originates from [20], where it is stated as an equivalence rather than an implication. But the converse is provable in our system.

The proof of Theorem 6 is the most difficult result of the paper, and we present it in full in the Appendix. The key difficulty of the proof consists in guaranteeing that the natural topology for which $\Box$ acts as interior operator is exactly the topology generated by the neighborhood family associated to $\Box_0$. Though the main steps of the proof involve known methods (a canonical quasi-model construction, a filtration argument, and then making multiple copies of the worlds), addressing the above-mentioned difficulty requires an innovative use of these methods, and a careful treatment of each of the steps. The proofs of the other results are standard, and so are left for the extended version of this paper, available at http://www.illc.uva.nl/Research/Publications/Reports/.

4 Further Developments and Future Work

The above-mentioned extended version contains an investigation of several types of evidential *dynamics* (building on the work in [21]), as well as complete axiomatizations of the corresponding dynamic-epistemic logics.

One line of further inquiry involves adding to the semantic structure a larger set $E_0^\diamond \supseteq E_0$ of *potential evidence*, meant to encompass all the evidence that might be learnt in the future. This would connect well with the topological program in Inductive Epistemology [10], based on a learning-theoretic investigation of convergence of beliefs to the truth in the limit, when the agent observes a stream of incoming evidence.

We also plan to extend our framework to notions of *group knowledge* for a group G. There are at least two different natural options for *common knowledge*: the Aumann concept (the infinite conjunction of "everybody knows that everybody knows etc"), and Lewis' concept, based on *shared evidence* (the intersection $\bigcap_{a \in G} E_0^a$ of the evidence families E_0^a of all agents $a \in G$). Similarly, there are now two different models for a *group's epistemic potential*: the standard concept of distributed knowledge, versus the one obtained by *sharing the evidence* (i.e. taking the union $E_0^G = \bigcup_{a \in G} E_0^a$ of all the evidence families E_0^a).

Acknowledgments We thank Johan van Benthem, Kevin Kelly, Thomas Icard and Eric Pacuit for feedback and discussions on the topic of this paper. We also thank the anonymous referees of WoLLIC 2016 for their useful comments. A. Özgün's work was funded by the European Research Council grant EPS 313360. Sonja Smets' research was funded by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement no. 283963.

Appendix: Proof of Theorem 6

A *quasi-model* is a tuple $\mathcal{M} = (X, E_0, \leq, V)$, where: $E_0 \subseteq \mathcal{P}(X)$ satisfies the same constraints as a topo-e-model, V is a valuation, $\leq$ is a preorder s.t. every $e \in E_0$ is upward-closed wrt $\leq$. The *semantics* is the same as on topo-e-models, except that $\Box$ gets a Kripke semantics: $\|\Box\phi\| := \{x \in X \mid \forall y \in X(x \leq y \Rightarrow y \in \|\phi\|)\}$.

A quasi-model $\mathcal{M} = (X, E_0, \leq, V)$ is called *Alexandroff* if the topology τ_E is Alexandroff and $\leq = \sqsubseteq_E$ is the specialization preorder. There is a natural *bijection B between Alexandroff quasi-models and Alexandroff topo-e-models*, given by putting, for any Alexandroff quasi-model $\mathcal{M} = (X, E_0, \leq, V)$, $B(\mathcal{M}) := (X, E_0, \tau_E, V)$. Moreover, $\mathcal{M}$ *and* $B(\mathcal{M})$ *satisfy the same formulas of $\mathcal{L}_{\forall\Box\Box_0}$ at the same points*. So Alexandroff quasi-models are just another presentation of Alexandroff models.

Proposition 5 *Let $\mathcal{M} = (X, E_0, \leq, V)$ be a quasi-model. The following are equivalent:*

1. *$\mathcal{M}$ is Alexandroff (hence, equivalent to an Alexandroff topo-e-model);*
2. *τ_E coincides with the family of all upward-closed sets (with respect to $\leq$);*
3. *for every $x \in X$, $\uparrow x$ is in τ_E.*

Proof. (1 $\Rightarrow$ 3) Suppose $\mathcal{M}$ is Alexandroff, i.e., τ_E is Alexandroff and $\leq = \sqsubseteq_E$. Let $x \in X$. Then we have: $\uparrow x = \{y \mid x \leq y\} = \{y \mid x \sqsubseteq_E y\} = \{y \mid \forall U \in \tau_E(x \in U \Rightarrow y \in U)\} = \bigcap\{U \in \tau_E \mid x \in U\}$. Since τ_E is an Alexandroff space, we have $\bigcap\{U \in \tau_E \mid x \in U\} \in \tau_E$, and hence $\uparrow x = \bigcap\{U \in \tau_E \mid x \in U\} \in \tau_E$.

(3 $\Rightarrow$ 2) Let $Up(X)$ be the set of all upward-closed subsets of X. It is easy to see that $\tau_E \subseteq Up(X)$ (since τ_E is generated by E_0 and every element of E_0 is upward-closed). Now let $A \in Up(X)$. Since A is upward-closed, we have $A = \bigcup\{\uparrow x \mid x \in A\}$. Then, by (3) (and τ_E being closed under arbitrary unions), we obtain $A \in \tau_E$.

(2 $\Rightarrow$ 1) Suppose (2) and let $\mathcal{A} \subseteq \tau_E$. By (2), every $U \in \mathcal{A}$ is upward-closed; hence, $\bigcap \mathcal{A}$ is upward-closed, so by (2) $\bigcap \mathcal{A} \in \tau_E$. This proves that τ_E is Alexandroff. (2) also implies that $\uparrow x$ is the least open neighbourhood of x in τ_E, i.e., $\uparrow x \subseteq U$, for all U such that $x \in U \in \tau_E$. Therefore, $\leq \subseteq \sqsubseteq_E$. For the other direction, suppose $x \sqsubseteq_E y$. This implies, in particular, $y \in \uparrow x$ (since $x \in \uparrow x \in \tau_E$), i.e., $x \leq y$.

The proof of Theorem 6 goes through *three steps*: (1) strong completeness for quasi-models; (2) finite quasi-model property; (3) every finite quasi-model is modally equivalent to a finite Alexandroff quasi-model (hence, to a topo-e-model).

Proposition 6 *(STEP 1) $L_{\forall\Box\Box_0}$ is sound and strongly complete for quasi-models.*

PROOF Soundness is easy. Completeness goes via a canonical quasi-model:

Lemma 1 *(*Lindenbaum Lemma*) Every consistent set of sentences in $\mathcal{L}_{\forall\Box\Box_0}$ can be extended to a maximally consistent one.*

Proof. Standard.

Let us now fix a consistent set of sentence Φ_0. Our goal is to construct a quasi-model for Φ_0. By Lemma 1, there exists a maximally consistent theory T_0 s. t, $\Phi_0 \subseteq T_0$. For any two maximally consistent theories T and S, we put: $T \sim S$ iff for all $\phi \in \mathcal{L}_{\forall\Box\Box_0}$: $((\forall\phi) \in T \Rightarrow \phi \in S)$; and $T \leq S$ iff for all $\phi \in \mathcal{L}_{\forall\Box\Box_0}$: $((\Box\phi) \in T \Rightarrow \phi \in S)$.

Canonical Quasi-Model for T_0. This is a structure $\mathcal{M} = (X, E_0, \leq, V)$, where: $X := \{T : T$ maximally consistent theory with $T \sim T_0\}$; $E_0 := \{\widehat{\Box_0\phi} : \phi \in \mathcal{L}_{\forall\Box\Box_0}$ with $(\exists\Box_0\phi) \in T_0\}$, where we used notation $\hat{\theta} := \{T \in X : \theta \in T\}$; $\leq$ is the restriction of the above preorder $\leq$ to X; and $V(p) := \hat{p}$. In the following, variables $T, S, \ldots$ range over X.

Lemma 2 *$\mathcal{M}$ is a quasi-model.*

Proof. Easy verification.

Lemma 3 *(*Existence Lemma for $\forall$*) $\widehat{\exists\varphi} \neq \emptyset$ iff $\hat{\varphi} \neq \emptyset$.*

Proof. Easy (along standard lines of the so-called Diamond Lemma for $\exists$).

Lemma 4 *(*Existence Lemma for $\Box$*) $T \in \widehat{\Diamond\varphi}$ iff $(\exists) S \in \hat{\varphi}$ s. t. $T \leq S$.*

Proof. Standard again.

Lemma 5 *(*Existence Lemma for $\Box_0$*) $T \in \widehat{\Box_0\varphi}$ iff $(\exists)\, e \in E_0$ s. t. $T \in e \subseteq \hat{\varphi}$.*

Proof. Left-to-right: Assume $T \in \widehat{\Box_0\varphi}$, i.e. $(\Box_0\varphi) \in T$. From $T \in X$ and $T \sim T_0$ we get $(\exists\Box_0\varphi) \in T_0$. Taking $e := \widehat{\Box_0\varphi}$, we get $e \in E_0$ and $T \in e$. To show that $e \subseteq \hat{\varphi}$, we use the theorem $\Box_0\varphi \to \varphi$, which implies that $\widehat{\Box_0\varphi} \subseteq \hat{\varphi}$, i.e. $e \subseteq \hat{\varphi}$.

Right-to-Left: Let $T \in X$ and $e \in E_0$, s.t. $T \in e \subseteq \hat{\varphi}$. Then $e = \widehat{\Box_0\theta}$ for some θ s.t. $(\exists\Box_0\theta) \in T_0$. So $T \in e = \widehat{\Box_0\theta} \subseteq \hat{\varphi}$. We now prove the following:

Claim: *The set $\Gamma := \{\Box_0\theta\} \cup \{\forall\psi : \forall\psi \in T\} \cup \{\neg\varphi\}$ is inconsistent.*

Proof of Claim: Suppose that $\Gamma \nvdash \bot$. By Lemma 1, there exists some $S \in X$ s. t. $\Gamma \subseteq S$. From $(\neg\varphi) \in S$ we get $S \notin \hat{\varphi}$ (by the consistency of S), and from $(\Box_0\theta) \in S$ we get $S \in \widehat{\Box_0\theta}$. So $S \in \widehat{\Box_0\theta} \setminus \hat{\varphi}$, contradicting $\widehat{\Box_0\theta} \subseteq \hat{\varphi}$.

Given the Claim, there exists a *finite* $\Gamma_0 \subseteq \Gamma$ with $\Gamma_0 \vdash \bot$. By the theorem $(\forall\psi_1 \wedge \ldots \forall\psi_n) \leftrightarrow \forall(\psi_1 \wedge \ldots \psi_n)$, we can assume that $\Gamma_0 = \{\Box_0\theta, \forall\psi, \neg\varphi\}$, for some ψ s. t. $(\forall\psi) \in T$. From $\Gamma_0 \vdash \bot$ we get the theorem $(\Box_0\theta \wedge \forall\psi) \to \varphi$. Using the Monotonicity Rule for $\Box_0$, the formula $\Box_0(\Box_0\theta \wedge \forall\psi) \to \Box_0\varphi$ is also a theorem. From the axiom $\Box_0\theta \to \Box_0\Box_0\theta$ and the Pullout Axiom, we get the theorem $(\Box_0\theta \wedge \forall\psi) \to \Box_0\varphi$. Since $(\Box_0\theta) \in T$ and $(\forall\psi) \in T$, it follows that $(\Box_0\varphi) \in T$, i.e. $T \in \widehat{\Box_0\varphi}$, as desired.

Lemma 6 (Truth Lemma) *For every formula $\phi \in \mathcal{L}_{\forall\Box\Box_0}$, we have: $||\phi||_{\mathcal{M}} = \hat{\phi}$.*

Proof. Standard proof by induction on the complexity of ϕ.

Consequence: $T_0 \models_{\mathcal{M}} \Phi_0$. This proves Step 1 (Proposition 6).

Theorem 7 *(STEP 2) The logic $\mathcal{L}_{\forall\Box\Box_0}$ has Strong Finite Quasi-Model Property.*

PROOF OF THEOREM 7: Let ϕ_0 be a consistent formula. By Step 1, take T_0 a maximal consistent theory s.t. $\phi_0 \in T_0$, and let $\mathcal{M} = (X, E_0, \leq, V)$ be the canonical quasi-model for T_0. We will use two facts about this model:

1. $||\varphi||_{\mathcal{M}} = \hat{\varphi}$, for all $\varphi \in \mathcal{L}_{\forall\Box\Box_0}$,
2. $E_0 = \{\widehat{\Box_0\varphi} : (\exists\Box_0\varphi) \in T_0\} = \{||\Box_0\varphi||_{\mathcal{M}} : (\exists\Box_0\varphi) \in T_0\}$.

Let Σ be a *finite* set such that: (1) $\phi_0 \in \Sigma$; (2) Σ is closed under subformulas; (3) if $(\Box_0\varphi) \in \Sigma$ then $(\Box\Box_0\varphi) \in \Sigma$; (4) Σ is closed under single negations; (5) $(\Box_0\top) \in \Sigma$. For $x, y \in X$, put: $x \equiv_\Sigma y$ iff $\forall\psi \in \Sigma(x \in ||\psi||_{\mathcal{M}} \iff y \in ||\psi||_{\mathcal{M}})$, and denote by $|x| := \{y \in X : x \equiv_\Sigma y\}$ the equivalence class of x modulo $\equiv_\Sigma$. Also, put $X^f := \{|x| : x \in X\}$, and more generally put $e^f := \{|x| : x \in e\}$ for every $e \in E_0$.

We now define a "*filtrated model*" $\mathcal{M}^f = (X^f, E_0^f, \leq^f, V^f)$, by taking: as set of worlds the set X^f (of equivalence classes) defined above; as for the rest, we put: $|x| \leq^f |y|$ iff for all $(\Box\psi) \in \Sigma : (x \in ||\Box\psi||_{\mathcal{M}} \Rightarrow y \in ||\Box\psi||_{\mathcal{M}})$; $E_0^f := \{e^f : e = \widehat{\Box_0\psi} = ||\Box_0\psi||_{\mathcal{M}} \in E_0$ for some ψ s. t. $(\Box_0\psi) \in \Sigma\}$; $V^f(p) := \{|x| : x \in V(p)\}$.

Lemma 7 *$\mathcal{M}^f$ is a finite quasi-model (of size bounded by a computable function of ϕ_0).*

Proof. X^f is finite, since Σ is finite so there are only finitely many equivalence classes modulo $\equiv_\Sigma$. In fact, the size is at most $2^{|\Sigma|}$. It's obvious that $\leq^f$ is a preorder, that $X^f \in E_0^f$ (since $X = ||\Box_0\top||_{\mathcal{M}}$ and $(\Box_0\top) \in \Sigma$, so $X^f \in E_0^f$) and that every $e^f \in E_0^f$ is non-empty (since it comes from some non-empty $e \in E_0$). So we only have to prove that the evidence sets are upward–closed: for this, let $e^f \in E_0^f$, with $e = \widehat{\Box_0\psi} \in E_0$, $(\Box_0\psi) \in \Sigma$ and let $|x| \in e^f$ and $|y| \in X^f$ s.t. $|x| \leq^f |y|$. We need to show that $|y| \in e^f$.

Since $|x| \in e^f$, there exists some $x' \equiv_\Sigma x$ s.t. $x' \in \widehat{\Box_0\psi} = ||\Box_0\psi||_{\mathcal{M}}$. From $(\Box_0\psi) \in \Sigma$ and $x' \equiv_\Sigma x$, we get $x \in ||\Box_0\psi||_{\mathcal{M}}$. By the theorem $\Box_0\psi \to \Box\Box_0\psi$, we have $x \in ||\Box\Box_0\psi||_{\mathcal{M}}$. But $(\Box\Box_0\psi) \in \Sigma$ (by the closure assumptions on Σ), so $|x| \leq^f |y|$ gives us $y \in ||\Box\Box_0\psi||_{\mathcal{M}}$. By the T-axiom $\Box\phi \to \phi$, we get $y \in ||\Box_0\psi||_{\mathcal{M}} = \widehat{\Box_0\psi} = e$, hence $|y| \in e^f$.

Lemma 8 (Filtration Lemma) *For every formula $\phi \in \Sigma$: $||\phi||_{\mathcal{M}^f} = \{|x| : x \in ||\phi||_{\mathcal{M}}\}$.*

Proof. Proof by induction on $\phi \in \Sigma$. The atomic case, inductive cases for propositional connectives and modalities $\forall\phi$ and $\Box\phi$ are treated as usual (-in the last case using the filtration property of $\leq^f$). We only prove here the inductive case for the modality $\Box_0\phi$:

Left-to-right inclusion: Let $|x| \in ||\Box_0\phi||_{\mathcal{M}^f}$. This means that there exists some $e^f \in E_0^f$ s.t. $|x| \in e^f \subseteq ||\phi||_{\mathcal{M}^f}$. By the definition of E_0^f, there exists some ψ s.t.: $(\Box_0\psi) \in \Sigma$ and $e = \widehat{\Box_0\psi} = ||\Box_0\psi||_{\mathcal{M}} \in E_0$. From $|x| \in e^f$, it follows that there is some $x' \equiv_\Sigma x$ s.t. $x' \in e = ||\Box_0\psi||_{\mathcal{M}}$, and since $(\Box_0\psi) \in \Sigma$, we have $x \in ||\Box_0\psi||_{\mathcal{M}} = e$. It is easy to see that we also have $e \subseteq ||\phi||_{\mathcal{M}}$. (Indeed, let $y \in e$ be *any* element of e; then $|y| \in e^f \subseteq ||\phi||_{\mathcal{M}^f}$, so $|y| \in ||\phi||_{\mathcal{M}^f}$, and by the induction hypothesis $y \in ||\phi||_{\mathcal{M}}$.) So we have found an evidence set $e \in E_0$ s.t. $x \in e \subseteq ||\phi||_{\mathcal{M}}$, i.e., shown that $x \in ||\Box_0\phi||_{\mathcal{M}}$.

Right-to-left inclusion: Let $x \in ||\Box_0\phi||_{\mathcal{M}}$, with $(\Box_0\phi) \in \Sigma$. It is easy to see that $(\exists\Box_0\phi) \in x$ (by the theorem $\Box_0\phi \to \exists\Box_0\phi$) and so also $(\exists\Box_0\phi) \in T_0$ (since $x \in X$ so $x \sim T_0$). This means that the set $e := \widehat{\Box_0\phi} = ||\Box_0\phi||_{\mathcal{M}} \in E_0$ is an evidence set in the canonical model, and since $(\Box_0\phi) \in \Sigma$, we conclude that $e^f \in E_0^f$ is an evidence set in the filtrated model. We obviously have $x \in e$, and so $|x| \in e^f$. By the (T) axiom, $e = ||\Box_0\phi||_{\mathcal{M}} \subseteq ||\phi||_{\mathcal{M}}$, and hence $e^f \subseteq \{|y| : y \in ||\phi||_{\mathcal{M}}\} = ||\phi||_{\mathcal{M}^f}$ (by the induction hypothesis). Thus, we have found $e^f \in E_0^f$ s.t. $|x| \in e^f \subseteq ||\phi||_{\mathcal{M}^f}$, i.e., shown that $|x| \in ||\Box_0\phi||_{\mathcal{M}^f}$.

Theorem 8 *(STEP 3) Every finite quasi-model is modally equivalent to a finite Alexandroff quasi-model (and so to a topo-e-model).*

PROOF OF THEOREM 8: Let $\mathcal{M} = (X, E_0, \leq, V)$ be a finite quasi-model. We form a new structure $\tilde{\mathcal{M}} = (\tilde{X}, \tilde{E_0}, \tilde{\leq}, \tilde{V})$, by putting: $\tilde{X} := X \times \{0, 1\}$; $\tilde{V}(p) := V(p) \times \{0, 1\}$; $(x, i)\tilde{\leq}(y, j)$ iff: $x \leq y$ and $i = j$; $\tilde{E_0} := \{e_i : e \in E_0, i \in \{0, 1\}\} \cup \{e_i^y : y \in e \in E_0, i \in \{0, 1\}\} \cup \{\tilde{X}\}$, where we used notations $e_i := e \times \{i\} = \{(x, i) : x \in e\}$ and $e_i^y := \uparrow y \times \{i\} \cup e \times \{1 - i\} = \{(x, i) : y \leq x\} \cup e_{1-i}$.

Lemma 9 *$\tilde{\mathcal{M}}$ is a (finite) quasi-model.*

Proof. Easy verification.

Notation: For any set $\tilde{Y} \subseteq \tilde{X}$, put $\tilde{Y}_X := \{y \in X : (y, i) \in \tilde{Y}$ for some $i \in \{0, 1\}\}$ for the set consisting of first components of all members of $\tilde{Y}$. It is easy to see that we have: $(\tilde{Y} \cup \tilde{Z})_X = \tilde{Y}_X \cup \tilde{Z}_X$, and $\tilde{X}_X = X$.

Lemma 10 *If $y \in e \in E_0$, $i \in \{0, 1\}$ and $\tilde{e} \in \{e_i, e_i^y\}$, then we have:*

1. $\tilde{e}_X = e$;
2. $e_i^y \cap e_i = \uparrow(y, i)$, *where* $\uparrow(y, i) = \{\tilde{x} \in \tilde{X} : (y, i)\tilde{\leq}\tilde{x}\} = \{(x, i) : y \leq x\}$.

Proof. 1. If $\tilde{e} = e_i$, then $\tilde{e}_X = (e \times \{i\})_X = e$. If $\tilde{e} = e_i^y$, then $\tilde{e}_X = (\uparrow y \times \{i\})_X \cup (e \times \{1 - i\})_X = \uparrow y \cup e = e$ (since e is upward-closed and $y \in e$, so $\uparrow y \subseteq e$).

2. $e_i^y \cap e_i = (\uparrow y \times \{i\} \cup e \times \{1 - i\}) \cap (e \times \{i\}) = (\uparrow y \cap e) \times \{i\} = \uparrow y \times \{i\} = \uparrow(y, i)$ (since $\uparrow y \subseteq e$).

Lemma 11 *$\tilde{\mathcal{M}}$ is an Alexandroff quasi-model (and thus also a topo-e-model).*

Proof. By Proposition 5, it is enough to show that, for every $(y, i) \in \tilde{X}$, the upward-closed set $\uparrow(y, i)$ is open in the topology τ_E generated by E_0. But this follows directly from part 2 of Lemma 10.

Lemma 12 (Modal-Equivalence Lemma) *For all $\varphi \in \mathcal{L}_{\forall\Box\Box_0}$: $||\varphi||_{\tilde{\mathcal{M}}} = ||\varphi||_{\mathcal{M}} \times \{0, 1\}$.*

Proof. Induction on φ. The base case, and the inductive steps for Boolean connectives and operators $\forall$ and $\Box$, are straightforward. We only prove the inductive step for $\Box_0$:

Left-to-Right Inclusion: Suppose that $(x, i) \in ||\Box_0\varphi||_{\tilde{\mathcal{M}}}$. Then there exists some $\tilde{e} \in \tilde{E}$ such that $(x, i) \in \tilde{e} \subseteq ||\varphi||_{\tilde{\mathcal{M}}} = ||\varphi||_{\mathcal{M}} \times \{0, 1\}$ (where we used the induction hypothesis for φ at the last step). From this, we obtain that $x \in \tilde{e}_X \subseteq (||\varphi||_{\mathcal{M}} \times \{0, 1\})_X = ||\varphi||_{\mathcal{M}}$. But by the construction of $\tilde{E}$, $\tilde{e} \in \tilde{E}$ means that either $\tilde{e} = \tilde{X}$ or there exist $e \in E_0$, $y \in e$ and $j \in \{0, 1\}$ such that $\tilde{e} \in \{e_j, e_j^y\}$. If the former is the case, we have $x \in \tilde{e}_X = X \subseteq ||\varphi||_{\mathcal{M}}$. Since $X \in E_0$, by the semantics of $\Box_0$, we obtain $x \in ||\Box_0\varphi||_{\mathcal{M}}$. If the latter is the case, by part 1 of Lemma 10, we have $\tilde{e}_X = e$, so we conclude that $x \in \tilde{e}_X = e \subseteq ||\varphi||_{\mathcal{M}}$. Therefore, again by the semantics of $\Box_0$, we have $x \in ||\Box_0\varphi||_{\mathcal{M}}$.

Right-to-Left Inclusion: Suppose that $x \in ||\Box_0\varphi||_{\mathcal{M}}$. Then there exists some $e \in E_0$ such that $x \in e \subseteq ||\varphi||_{\mathcal{M}}$. Take now the set $e_i = e \times \{i\} \in \tilde{E}$. Clearly, we have $(x, i) \in e_i \subseteq ||\varphi||_{\mathcal{M}} \times \{i\} \subseteq ||\varphi||_{\mathcal{M}} \times \{0, 1\} = ||\varphi||_{\tilde{\mathcal{M}}}$ (where we used the induction hypothesis for φ at the last step), i.e. we have $(x, i) \in ||\Box_0\varphi||_{\tilde{\mathcal{M}}}$.

Theorem 8 follows immediately from the above Lemma: the same formulas are satisfied at x in $\mathcal{M}$ as at (x, i) in $\tilde{\mathcal{M}}$. Theorem 6 is an immediate corollary of Theorem 8.

References

1. Baltag, A., Bezhanishvili, N., Özgün, A., and Smets, S. (2013) The topology of belief, belief revision and defeasible knowledge. *In Proc. of LORI 2013*, pp. 27–40, Springer, Heidelberg.
2. Baltag, A., Bezhanishvili, N., Özgün, A., and Smets, S. (2015) The topological theory of belief. *Submitted. http://www.illc.uva.nl/Research/Publications/Reports/PP-2015-18.text.pdf*.
3. Baltag, A. and Smets, S. (2006) Conditional doxastic models: A qualitative approach to dynamic belief revision. *In Proc. of WOLLIC 2006*, vol. 165, pp. 5–21.
4. Baltag, A. and Smets, S. (2008) A qualitative theory of dynamic interactive belief revision. *Texts in Logic and Games*, **3**, 9–58.
5. Clark, M. (1963) Knowledge and grounds: a comment on Mr. Gettier's paper. *Analysis*, **24**, 46–48.
6. Foley, R. (2012) *When is true belief knowledge?*. Princeton University Press.
7. Gettier, E. (1963) Is justified true belief knowledge? *Analysis*, **23**, 121–123.
8. Goranko, V. and Passy, S. (1992) Using the universal modality: Gains and questions. *J. Log. Comput.*, **2**, 5–30.
9. Grove, A. (1988) Two modellings for theory change. *Jour. of Phil. Logic*, **17**, 157–170.
10. Kelly, K. (1996) *The Logic of Reliable Inquiry*. Oxford University Press, Oxford.
11. Klein, P. (1971) A proposed definition of propositional knowledge. *Journal of Philosophy*, **68**, 471–482.
12. Klein, P. (1981) *Certainty, a Refutation of Scepticism*. University of Minneapolis Press.
13. Lehrer, K. (1990) *Theory of Knowledge*. Routledge.
14. Lehrer, K. and Paxson, T. J. (1969) Knowledge: Undefeated justified true belief. *Journal of Philosophy*, **66**, 225–237.
15. McKinsey, J. C. C. and Tarski, A. (1944) The algebra of topology. *Ann. of Math. (2)*, **45**, 141–191.
16. Özgün, A. (2013) *Topological Models for Belief and Belief Revision*. Master's thesis, University of Amsterdam, Amsterdam, The Netherlands.
17. Rott, H. (2004) Stability, strength and sensitivity: Converting belief into knowledge. *Erkenntnis*, **61**, 469–493.
18. Stalnaker, R. (2006) On logics of knowledge and belief. *Phil. Studies*, **128**, 169–199.
19. van Benthem, J. (2007) Dynamic logic for belief revision. *JANCL*, **17**, 129–155.
20. van Benthem, J., Fernández-Duque, D., and Pacuit, E. (2014) Evidence and plausibility in neighborhood structures. *Annals of Pure and Applied Logic*, **165**, 106–133.
21. van Benthem, J. and Pacuit, E. (2011) Dynamic logics of evidence-based beliefs. *Studia Logica*, **99(1)**, 61–92.